



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Is dishonesty persistent?

Citation for published version:

Belot, M & Van De Ven, J 2019, 'Is dishonesty persistent?', *Journal of Behavioral and Experimental Economics*, vol. 83, 101464. <https://doi.org/10.1016/j.socec.2019.101464>

Digital Object Identifier (DOI):

[10.1016/j.socec.2019.101464](https://doi.org/10.1016/j.socec.2019.101464)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of Behavioral and Experimental Economics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Is Dishonesty Persistent?

Michèle Belot and Jeroen van de Ven^{*}

This version: September 4, 2019

Abstract: We study if (dis)honest behavior is persistent. We investigate this by exposing participants to different incentives to lie over time. Some participants are first exposed to high incentives and then to lower incentives; for others the reverse. If (dis)honest behavior is persistent, the propensity to lie depends on past incentives. We find no evidence of persistence in honest or dishonest behavior. Exposing participants first to high incentives does not result in a lasting positive effect on dishonesty after the incentives are lowered away. Similarly, after correcting for a time trend, subjects still respond strongly to high incentives after facing low incentives.

Keywords: lying, dishonesty, persistence, habits, incentives, experiment

^{*} Belot: European University Institute, Florence, and School of Economics, University of Edinburgh. Michele.Belot@EUI.eu. Van de Ven: Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute, j.vandeven@uva.nl. Special thanks go Daniel Chen, Gönül Doğan, Uri Gneezy, Aaron Kamm, Emily Ledebor, Anne van Lieshout, and Aljaz Ule, the Associate Editor (Bradley Ruffle) and two anonymous referees for their valuable suggestions and/or help with running the experiments. We also thank participants at seminars and conferences at NYU Abu Dhabi, Higher School of Economics (Moscow), The Human Face in Economics workshop (Toulouse), Experimental Methods Conference (Ixtapa), GATE (Lyon), UCSD (Rady School of Management), EUR (Rotterdam), EASP (Amsterdam), and the Collier School of Management workshop (Tel Aviv). Financial support from the Research Priority Area Behavioral Economics (UvA) is gratefully acknowledged.

1. Introduction

In many markets with informational asymmetries, the incentives to mislead customers are high, such as in markets for financial products, car and electronics repairs, and medical services (Dulleck et al., 2011; Kerschbamer, 2016). It has been well documented that people are prepared to deceive others, and that their willingness to lie responds strongly to incentives (Gneezy, 2005; Sutter, 2009; Erat and Gneezy, 2012)¹.

In this paper, we are interested in the *persistence* of honest or dishonest behavior. To our knowledge, there is only little known about the persistency of dishonesty. There are studies on dishonesty that involve multiple rounds where participants decide whether to lie or not. Abeler et. (2019) reports on 90 experimental studies in the literature using the Fischbacher and Fölmi-Heusi (2013) paradigm whereby people have to report the outcome of a random event (such as a coin toss or a die roll) they observe privately and where the report determines their payoff. They report that in experiments with repeated decisions there is no clear trend, i.e. one does not observe a tendency to be more or less truthful over time. It is, however, not possible to establish (non-) persistency from studying time trends alone, without inducing some exogenous shock to dishonesty. A number of other recent studies suggest there may be a slippery slope in dishonest and unethical behavior (Welsh et al., 2015; Garrett et al., 2016). The idea is that small benefits to lie can induce some dishonesty, and when these benefits are gradually increased this ultimately results in more dishonesty than if the benefits increase abruptly.

In this study, we study persistence by manipulating the incentives to lie over time. We first expose experimental participants to high or low incentives, and then reverse the incentives. We hypothesize that dishonesty may be persistent, that is, that the propensity to lie under low incentives is higher when participants have first been exposed to high incentives. There are several reasons why we could expect this. First, dishonesty, as many other behaviors, might become habitual over time (Fehr and Engelmann, 2016). Secondly, some people might be willing to lie when the incentives are high, but still feel some discomfort. In an attempt to justify their behavior and preserve a good self-image, they may downplay any wrongdoing (“it’s just a game”).

¹ Kajackaite and Gneezy (2017) do not find strong effects of incentives on behavior in nonstrategic settings.

Such motivated beliefs are supported by empirical evidence.² In both cases, dishonesty would persist even if incentives are reduced or taken away. On the other hand, there are also reasons to believe that dishonesty may not be persistent. People may engage in ‘conscience accounting’ (e.g., Gneezy et al., 2014b) or ‘moral licensing’ (Merritt et al., 2010), and try to compensate past immoral behavior by good deeds. The flip side of this prediction is that honesty may be persistent as well. If the incentives to lie are initially low, people may refrain from lying later on, even if there are higher incentives to do so. Again, this could be due to habit formation or motivated beliefs.

We report results from two laboratory experiments. In both experiments, participants are asked to play a sender-receiver game, in which senders have private information. The game is framed as an exchange between sellers (the senders) and buyers (the receivers). The sellers have private information about the quality of the product they are selling. The buyers’ task is to guess the quality of the product. It is in the self-interest of buyers to guess the quality correctly. Sellers, on the other hand, gain from convincing buyers that the product is of high quality even if it is not. Sellers and buyers can briefly communicate with each other before buyers make their guesses.

Participants are asked to play this game for several rounds. We change the size of the incentives halfway through the experiment. One group starts with large incentives to lie and the other group starts with small incentives. We reverse the incentives after several rounds, and examine how behavior is affected by past incentives.

The experiments vary in context and sample. Experiment 1 involves students from a class in experimental economics and communication between buyers and sellers is face-to-face. Experiment 2 involves participants from a larger subject pool, and the experiment was computerized. Participants could communicate using chat messages.

In both experiments, we find a large drop in lying rates after reducing the incentives, and only a small increase in lying rates after increasing the incentives. Thus, it is not the case that the group starting with high stakes remains more prone to lying when the stakes are reduced. In fact,

² This literature goes back to at least Festinger (1957). Some recent work in the realm of unethical behavior includes Konow (2005), Gneezy et al. (2014a), and see Bénabou and Tirole (2006).

this group shows the lowest lying rate after reducing the stakes. In Experiment 2, the group starting with low incentives to lie already exhibits a high lying rate when facing low incentives, and so the modesty of the increase in lying rates could be driven by a ceiling effect. In an additional treatment, we do indeed find evidence for this.

The large drop in lying rates when going from high to low incentives suggests that this group could even be ‘overcompensating’ for their past behavior. However, this pattern could also be driven by a time trend. To test this, Experiment 2 includes an additional treatment in which participants face low incentives in all rounds. We do indeed observe a decrease in lying rates over time for that group. The overcompensation effect remains, but it is no longer statistically significant after correcting for a general time trend. This time trend can also explain the modest increase in lying when stakes are increased: subjects still respond strongly to incentives but the effect is partly offset by the negative time trend.

One interesting lesson from these studies is that manipulating incentives appears to be very effective at influencing dishonesty. We find no evidence that dishonesty or honesty are persistent.

The paper contributes to the large and growing literature on dishonesty and deception. As mentioned earlier, a couple of studies using different experimental paradigms (Welsh et al., 2015; Garrett et al., 2016) find evidence for increasing dishonest behavior over time. Welsh et al. (2015) uses a design that shares some similarity with ours. In their first study, they implement two treatments with increasing rewards for cheating over different rounds. In one of the treatments this increase in rewards is abrupt, in the other treatment it is more gradual. Interestingly, even though in the last round the rewards are equal in both treatments, they find more cheating behavior when the reward increased gradually. Their task is different from ours, with no strategic interaction. They do not have a treatment in which incentives decrease, so they cannot study what happens to cheating behavior if incentives are removed. Nevertheless, their study shows that when the incentives change abruptly from low to high, the propensity to lie is not lower in the high condition than when the incentives have been high all the way through, that is, they find no evidence of persistence in honesty. In a coin-tossing task, Ruffle and Wilson

(2018) also vary the stakes within subjects, but subjects know in advance that the incentives will increase, so that they can in principle time their honesty strategically. Garrett et al. (2016) do not expose the same subjects to different incentives. Other studies that examine the persistence of lying are typically non-experimental in nature, so that no causal claims can be made, and they tend to rely on self-reported dishonesty (e.g., Knopp et al., 2017).

Another related literature studies long-term effects of incentives in other domains. For instance, Charness and Gneezy (2009) show that incentives can be used to induce habits. Paying people to exercise had positive effects on gym attendance even after incentives were withdrawn. Falk et al. (2006) show persistent effects of a minimum wage in a laboratory experiment. Other studies show evidence of crowding-out effects. In those cases, incentives have long-term negative effects after they are removed (e.g., Deci, 1971; Deci and Ryan, 1975; Gneezy and Rustichini, 2000; Frey and Jegen, 2001).

More generally, order effects have been studied in a variety of contexts, often to control for experience. Many studies do not find evidence of order effects, but some do (e.g., Owens and Kagel, 2010). As far as we know, order effects have not been studied in the domain of dishonesty. Since the cause of order effects can be very different (e.g., habits, self-justification), we believe it is interesting to study this in the context of dishonesty.

2. Experiment 1

2.1. Experimental setup and procedures

The experimental setup closely follows that in Belot and Van de Ven (2017). Participants were randomly allocated the role of buyer or seller. Every buyer was paired with a seller. Each seller received a description of a hypothetical product that they were to sell. This product could be in good condition (high quality) or in bad condition (low quality). A random draw from a deck determined the state of the product. The deck contained 5 green and 5 red cards, and after every draw the card was placed back in the deck. If the card drawn was green, the product was in good condition. If it was red, the product was in bad condition. The colour of the card remained private information to the seller (and the experimenter) and was never revealed to the buyer.

After drawing the card, the seller and buyer interacted face-to-face. In the first 10 seconds, the seller could make a claim about the quality of the product (high or low). In the following 90 seconds, buyers and sellers could chat free format. After that, sellers were asked to privately write down the claim they made about their product. Buyers were also asked to privately write down the seller's claim, and to make a guess about the quality of the product. We also asked sellers how confident they felt that buyers would guess green, and buyers were asked how confident they felt that the card was green, both on a 11-point scale.

We varied the sellers' payoffs (see Table 1). In the **high incentives** condition, a seller would earn €0 if the buyer guessed low quality and €30 if the buyer guessed high quality. In the **low incentives** condition, a seller would earn €12 if the buyer guessed low quality and €18 if the buyer guessed high quality. Note that in both cases the sellers' payoffs were independent of the actual quality. Buyers always earned €20 for a correct guess and €0 for an incorrect guess. Thus, in terms of monetary payoffs, sellers are always better off if they can convince the buyers that they drew a green card, while the buyers are better off guessing the actual colour of the seller's card. The payoff structure was always made common knowledge.

Table 1: Payoff structure

Seller's card (random draw)	Buyer's guess	Seller's earnings		Buyer's earnings
		Low	High	
		incentives	incentives	
Red	Red	12	0	20
Red	Green	18	30	0
Green	Red	12	0	0
Green	Green	18	30	20

The treatment variation was the order in which they were exposed to the different conditions (see Table 2). There were two parts, with 6 to 8 rounds in each part (depending on the

group size). One group started in the low incentives condition, while the other group started in the high incentives condition. In part two, the conditions were reversed between the groups. We refer to these groups as *High-Low* and *Low-High*.

Table 2: Treatments Experiment 1

Treatment	Incentives		N
	Part 1	Part 2	
Low-High	Low	High	70
High-Low	High	Low	68

Participants kept their role throughout the experiment. Sellers kept the same product description, but drew a new card at the start of every round. They were re-matched every round, and never met the same buyer twice within the same part. They did not receive any feedback between rounds. At the end, one round was randomly selected for payment. To preserve anonymity, we did not reveal which round was selected.

Sessions were run in 2013 and 2014 in the CREED lab (University of Amsterdam) as part of a course in experimental economics. There were 138 participants in total (47 percent female, mean age 24), divided over eight sessions. The experiment was run with paper and pencil. Instructions were read aloud after which participants were given time to read the instructions themselves. Instructions for part 2 were distributed at the end of part 1. Each session ended with a short survey to collect some basic demographics³ and participants were asked to indicate for each paired participant whether they were acquainted, friends, or did not know each other. A session lasted around 90 minutes and average earnings were €15.58. No show up fee was paid because attendance was compulsory.

³ See the Appendix for the survey questions. These questions were collected for another project and are not used in this study.

2.2. Hypotheses

Before we formulate our hypotheses we first define our measure of persistence. Let the mean lying rate in part 1 be given by:

$$y_1(h_1) = a + b \cdot h_1.$$

Here, $a \geq 0$ is some positive constant and $h_i \in \{0,1\}$ where $h_i = 1$ corresponds to high incentives. The parameter b measures the responsiveness to incentives and we assume it is positive. In part 2, let the mean lying rate be given by:

$$y_2(h_1, h_2) = a + (1 - \lambda)b \cdot h_2 + \lambda b \cdot h_1.$$

The parameter $\lambda \in [0,1]$ is our measure of persistency. If $\lambda = 0$, past incentives have no effect on behavior in part 2 (no persistence). If $\lambda = 1$, behavior in part 2 depends on past incentives but not on current incentives (complete persistence). For $\lambda \in (0,1)$, there is partial persistency. In the remainder, we speak of persistency whenever $\lambda > 0$.

We now briefly formulate the main hypotheses we will test.

Null-hypothesis: The null hypothesis is that (dis)honesty is not persistent (i.e., $\lambda = 0$). In particular, the propensity to lie does not depend on incentives that individuals faced in the past.

Alternative hypotheses: The alternative hypothesis is that past behavior persists, such that $\lambda > 0$. In that case, the propensity to lie under the same incentives should depend on the past incentive structure. It could be that the degree of persistence depends on the initial incentives, that is, it could be that the psychological mechanism underlying persistence in dishonest behavior differs from the mechanism underlying persistence in honest behavior. We will therefore test two hypotheses: (1) the propensity to lie when facing low incentives is higher in the high-low condition than in the low-high condition, so that $y_2(1,0) > y_1(0)$, and (2) the propensity to lie when facing high incentives is lower in the low-high condition than in the high-low condition, so that $y_2(0,1) < y_1(1)$.

In this Experiment, we ignore possible time trends in dishonesty. We will consider the possibility of time trend in Experiment 2.

2.3. Results

Figure 1 shows the proportion of sellers that lie, i.e., those with a red card who claim to have a green card.⁴ Combined over the treatments, we find a strong effect of incentives: sellers lie 71 percent of the time when incentives are high versus 38 percent when incentives are low ($p < 0.001$, signed rank test⁵). Second, we find that the responsiveness to a change in incentives differs strongly across the two groups. The higher incentives increase the lying rate by 60 percentage points in *High-Low* but only by 6 percentage points in *Low-High*. This difference (60 versus 6) is highly significant ($p < 0.001$, two-sided Mann-Whitney test).

The lying rate in *High-Low* shows a large drop after lowering the incentives, from 87 to 27 percent ($p < 0.001$, signed rank test). Thus, those who lied in part 1 when incentives were high largely refrain from lying when the incentives become low. There is no persistency in dishonesty, i.e., $y_2(1,0) \neq y_1(0)$. Thus, we do not find that $\lambda > 0$. In fact, if anything it seems that $\lambda < 0$, suggesting that participants overcompensate: the lying rate when facing low incentives is lower in the *High-Low* treatment than in the *Low-High* treatment (the latter is 48 percent), and the difference is significant ($p = 0.024$, two-sided Mann-Whitney test).

The lying rate in *Low-High* hardly increases with higher incentives, and the difference is not significant ($p = 0.392$, signed rank test). Comparing lying rates when the incentives are high, we find that those in *Low-High* end up lying less than those in *High-Low* (54 versus 87 percent, $p = 0.002$, two-sided Mann-Whitney test). There appears to be persistency in honesty: $y_2(0,1) < y_1(1)$. However, this can be driven by a negative time-trend. This is something that we explore in the Experiment 2.

⁴ We disregard lying by sellers with a red card. This strategy is very rarely used (less than 3 percent of the time).

⁵ To be conservative, here and elsewhere we take the mean of a participant over all rounds (within a part, if relevant) as the unit of observation. Note however that the results of the statistical tests need to be interpreted with caution, as observations are not independent within a session.

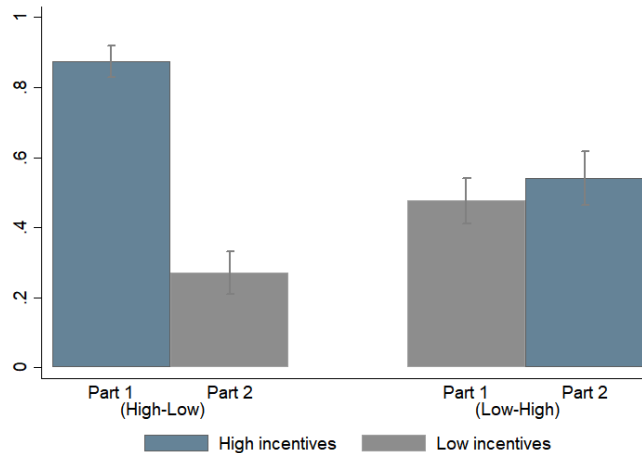


Figure 1: Lying rates by treatment and part. Proportion of sellers with a red card that claims the card is green. Error bars indicate ± 1 s.e.

3. Experiment 2

We replicated the experiment in a different experimental environment, with a larger sample of participants and with an additional treatment where the incentives to lie remain unchanged, to identify a possible time trend in honesty.

3.1. Experimental setup and procedures

The basic structure of Experiment 2 is the same as that of Experiment 1. This time, we recruited participants from the CREED database, instead of students from the experimental economics class. We replaced the face-to-face communication with chat messaging. To keep it personal, all participants took a selfie prior to starting, and these pictures were shown to the paired participants. Figure A2-1 in Appendix 2 shows a screenshot of the chat environment. To keep things simple, we did not use the hypothetical product descriptions, but only referred to

red and green cards.⁶ The payoffs were identical to those in Experiment 1, with the minor change that buyers earned €30 for a correct guess instead of €20.

To address the possible concern of time trends, we added an additional treatment, *Low-Low*, where participants face low incentives in both parts of the experiment (see Table 3).

Table 3: Treatments Experiment 2

Treatment	Incentives		N
	Part 1	Part 2	
Low-High	Low	High	64
High-Low	High	Low	60
Low-Low	Low	Low	64

Sessions were run in 2015 in the CREED lab (University of Amsterdam). There were 188 participants in total (46 percent female, mean age 22), divided over 12 sessions. There were always 16 participants per session, except for one session which had 12 participants due to a low show-up rate. In each round, half of the sellers saw a red card on their screen and the other half a green card. Which sellers received the red/green card was randomized in each round. The experiment was computerized and programmed in Php/MySQL. Instructions were given on the screen. A hand-out was distributed with the summary of the instructions. Each session ended with a short survey and participants were asked to indicate for each paired participant whether they were acquainted, friends, or did not know each other. A session lasted around 80 minutes and average earnings were €22.24, including a fixed show-up fee of €5.

⁶ In terms of spotting deception, Belot and Van de Ven (2017) show that it doesn't matter much whether the context is rich (with product quality associated to green and red cards) or abstract (with green and red cards).

3.2. Hypotheses

We test the same hypotheses as in Experiment 1, but Experiment 2 introduces an additional treatment to control for a possible time trend in dishonesty.

In that case, the mean lying rate in part 2 is given by:

$$\hat{y}_2(h_1, h_2) = a + (1 - \lambda)b \cdot h_2 + \lambda b \cdot h_1 + \phi,$$

where the parameter ϕ captures the time trend, and it can be positive or negative. We implicitly assume the time trend to be independent of h_1 . In this case and ignoring possible floor or ceiling effects, persistence in dishonesty implies $y_2(1,0) - \phi > y_1(0)$ and persistence in honesty implies $y_2(0,1) - \phi < y_1(1)$.

3.3. Results

Figure 2 shows the results. Several of the patterns look very similar to those we found in Experiment 1. In particular, in treatment *High-Low* we see a substantial drop in lying rates when moving from high to low incentives. In treatment *Low-High* the lying rate increases with incentives, where, as before, the increase is modest.

We also again observe overcompensation: lying rates are lowest among participants who face low incentives after first being exposed to high incentives. The difference with participants who started with low incentives is 36 percentage points (37 versus 73 percent). However, as can be seen from lying rates in treatment *Low-Low*, this is at least partially attributable to a downward trend that is present even when incentives are held constant (corresponding to a negative value of ϕ). The lying rate in part 2 in *High-Low* (37 percent) is lower than the lying rate observed in part 2 in *Low-Low* (49). However, this difference is not significant, and we cannot reject that $y_2(1,0) = y_2(0,0)$ ($p = 0.182$, two-sided Mann-Whitney test).⁷

It is also worth pointing out that the lying rates in treatment *Low-High* are substantially higher than in Experiment 1. Thus, there is not much honest behavior in the low-stake condition

⁷ This might be an underestimate. Note that the lying rate in part 1 of *Low-Low* is already a bit below that in *Low-High*. If they would have started at the same level in part 1 as in *Low-High* the effect would be larger. Of course, we have no way of determining which base level is more accurate.

in Experiment 2 to start with. We speculate that the higher degree of anonymity in Experiment 2 compared to Experiment 1 is driving this. Participants in Experiment 2 are less likely to know each other and meet each other again in the future⁸, and chat messages might be perceived as more anonymous than face-to-face communication. Whatever is the cause, it means that we have to be cautious in concluding that participants in this treatment are not very responsive to incentives; with the high base rate, this could be due to a ceiling effect.⁹

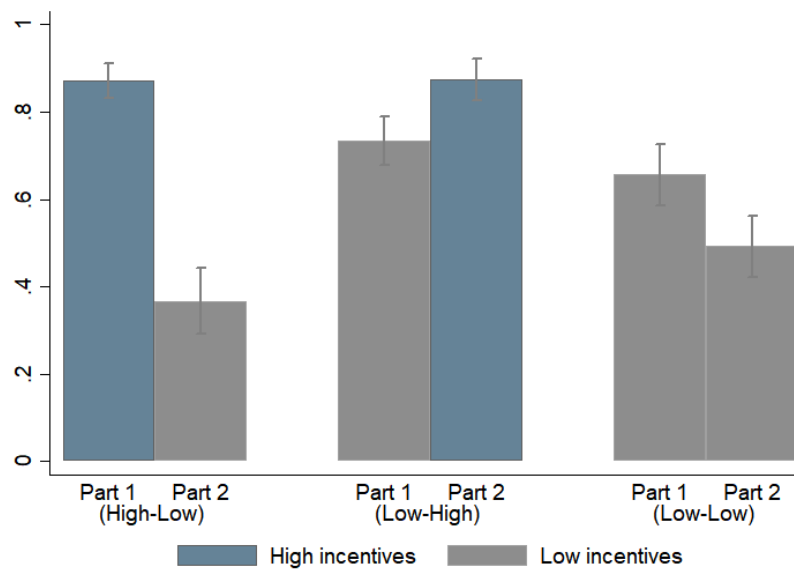


Figure 2: Lying rates by treatment and part. Proportion of sellers with a red card that claims the card is green. Error bars indicate ± 1 s.e.

3.4. Alternative explanations?

⁸ In Experiment 1, 22 percent of participants indicated to be acquainted and 17 percent indicated to be friends. In Experiment 2, only 4 percent is acquainted and 4 percent is friends with the paired participant. We do not find a clear effect of friendship on the lying rate. In Experiment 1, the overall lying rate is 57 percent when participants don't know the paired participant, 45 percent when they are acquainted, and 56 percent when they are friends. In experiment 2, the lying rates are 67 (don't know), 68 (acquainted), and 49 (friends).

⁹ We indeed find some evidence for a ceiling effect. We ran two additional sessions ($N = 28$) in which we decreased the stakes in part 1 for sellers even further. Sellers could gain either €14 or €16, depending on whether the buyer guessed the card was red or green. Even for those very low incentives, 54 percent of sellers with a red card lied about the color, and this increased to 86 percent when they faced high incentives.

An important question is whether other factors than the incentives themselves could have been driving the sharp fall in lying rates going from high to low incentives. We briefly discuss three candidates: tiredness, lower credibility, and conscience accounting.

One possibility is that participants get tired of lying. This would be consistent with the overall downward trend that we find in treatment *Low-Low*. While we cannot rule this out completely, we do not believe that this is the major force. Figure 3 shows the mean lying rate over rounds. If tiredness played a role, we would expect to see a gradual decline in lying rates. What we observe, however, is a sudden drop after the change in incentives and, if anything, an increase in the rate of lying in Part 1. This is the case in both experiments. Insofar as the length of the chat message is an indicator of effort, we also do not observe signs of tiredness. The number of words by sellers steadily increases over the rounds in part 1 of *High-Low*, from about 31 to 41 words per seller. There is also no difference in length between the two parts; sellers write about 37 words on average per round in both parts. More generally, it seems that sellers with a green card are the ones that exert the most effort in persuading the buyer: on average they write about 42 words per round. Thus, if anything, honesty could be more tiring.

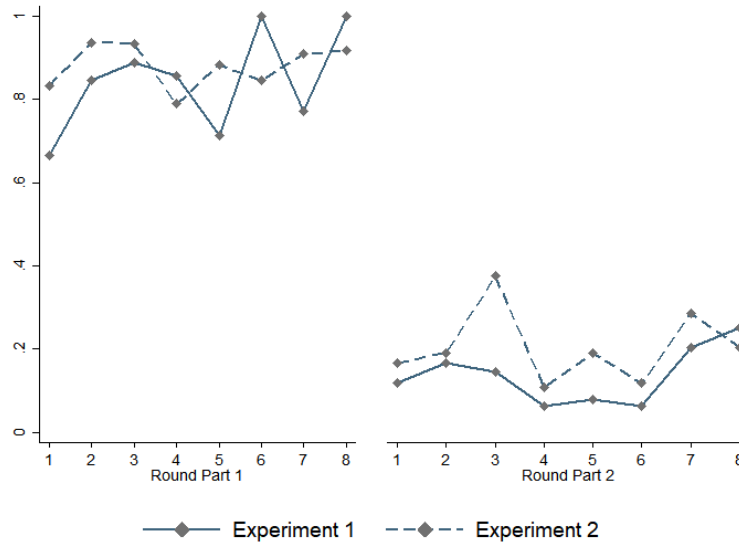


Figure 3: Lying rates in “High-Low” over time. Proportion of sellers in treatment *High-Low* with a red card that claims the card is green over the rounds.

Another possibility is that sellers anticipate a lower credibility of claims to have a green card in part two. With almost all sellers claiming to have a green card in part one, buyers naturally become more suspicious.¹⁰ Sellers may not be willing to lie if they think they won’t be believed in any case. We do not find evidence to support this. When the incentives are low and the seller claims to have a green card, buyers are even a bit more confident in the seller’s truthfulness in treatment High-Low than in treatments where incentives were low in part 1, though the difference between treatments is never significant ($p > 0.166$ in both experiments, two sided Mann-Whitney test). Furthermore, in the face of the sudden drop, one would expect this effect to be already present at the start of part 2. However, comparing confidence in part 1 to the first two rounds in part two, we find very similar effects.¹¹

¹⁰ In this respect, a cleaner design would be to have a new set of inexperienced sellers in part 2. We thank a referee for pointing this out.

¹¹ We also find that when incentives are low, sellers that lie in treatment High-low are a bit more confident that the buyer believes them compared to sellers that faced low incentives in part 1 and lied. Note, however, that selection effects make these results hard to interpret, since out of those with a red card only the sellers that are confident may want to lie.

Finally, it could be that people want to do good in part 2 to offset the unethical choices in part 1, something called “conscience accounting” (Gneezy et al., 2014b). Also, this explanation does not seem in line with the sudden drop between the parts, however.¹²

4. Conclusion

Our main finding is that dishonesty responds strongly to incentives and does not exhibit persistence: participants first exposed to high incentives to lie do not lie more when facing low incentives than those who were never exposed to large incentives. Similarly, participants first exposed to low incentives to lie do not lie less when facing high incentives than those who start with high incentives. We also find evidence for a negative trend in dishonesty when incentives are low in all rounds, that is, participants tend to lie less over time.

We conclude that the adage “Once a cheater always a cheater” does not apply here. Of course, that does not mean that lying will never be persistent. Habit formation takes time. Longer repeated exposure to large incentives may result in more persistent dishonesty. Testing this would be an exciting avenue for further research in our view.

¹² All or some of the factors discussed above (tiredness, lower credibility, conscience accounting) could be responsible for the downward trend in *Low-Low*, which does not show a sharp drop between parts.

5. References

- Abeler, J., Nosenzo, D., and Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *Forthcoming*.
- Belot, M., & Van de Ven, J. (2017). How private is private information? The ability to spot deception in an economic game. *Experimental Economics*, 20(1), 19-43.
- Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American economic review*, 96(5), 1652-1678.
- Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, 77(3), 909-931.
- Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, 18(1), 105.
- Deci, E. L., & Ryan, R. M. (1975). *Intrinsic motivation*. John Wiley & Sons, Inc.
- Dulleck, U., Kerschbamer, R., & Sutter, M. (2011). The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *American Economic Review*, 101(2), 526-55.
- Engelmann, J. B., & Fehr, E. (2016). The slippery slope of dishonesty. *Nature neuroscience*, 19(12), 1543.
- Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, 58(4), 723-733.
- Falk, A., Fehr, E., & Zehnder, C. (2006). Fairness perceptions and reservation wages—the behavioral effects of minimum wage laws. *The Quarterly Journal of Economics*, 121(4), 1347-1381.
- Festinger, L. (1957). A theory of cognitive dissonance. Stanford University Press.
- Fischbacher, Urs and Franziska Föllmi-Heusi (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association* 11 (3), 525–547.
- Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of economic surveys*, 15(5), 589-611.
- Garrett, N., Lazzaro, S. C., Ariely, D., & Sharot, T. (2016). The brain adapts to dishonesty. *Nature neuroscience*, 19(12), 1727.

- Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1), 384-394.
- Gneezy U., Saccardo S., Serra-garcia M. and Van Veldhuizen R. (2014a). Motivated Self-Deception and Unethical Behavior, UCSD mimeo.
- Gneezy, U., Imas, A., & Madarász, K. (2014b). Conscience accounting: Emotion dynamics and social behavior. *Management Science*, 60(11), 2645-2658.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29(1), 1-17.
- Kajackaite, A., & Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior*, 102, 433-444.
- Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American economic review*, 90(4), 1072-1091.
- Knopp, K., Scott, S., Ritchie, L., Rhoades, G. K., Markman, H. J., & Stanley, S. M. (2017). Once a Cheater, Always a Cheater? Serial Infidelity Across Subsequent Relationships. *Archives of sexual behavior*, 46(8), 2301-2311.
- Kerschbamer, R., Neururer, D., & Sutter, M. (2016). Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proceedings of the National Academy of Sciences*, 113(27), 7454-7458.
- Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and personality psychology compass*, 4(5), 344-357.
- Owens, M. F., & Kagel, J. H. (2010). Minimum wage restrictions and employee effort in incomplete labor markets: An experimental investigation. *Journal of Economic Behavior & Organization*, 73(3), 317-326.
- Ruffle, B. J., & Wilson, A. E. (2018). The truth about tattoos. *Economics Letters*, 172, 143-147.
- Sutter, M. (2009). Deception through telling the truth?! Experimental evidence from individuals and teams. *The Economic Journal*, 119(534), 47-60.
- Welsh, D.T., Ordóñez, L.D., Snyder, D.G. & Christian, M.S. (2015). The slippery slope: how small ethical transgressions pave the way for larger future transgressions, *Journal of Applied Psychology* 100, 114–127.

Appendix A: Instructions

Below are the instructions and survey questions used in Experiment 2 for the treatment in which they start with high incentives. The other instructions are available upon request.

[Page 1]

Instructions Part 1

Welcome to this experiment. Please read the instructions carefully. We ask you to remain silent throughout the experiment.

There are two types of roles in this experiment: **sellers** and **buyers**. We will randomly assign a role to you and you will keep that role throughout the entire experiment.

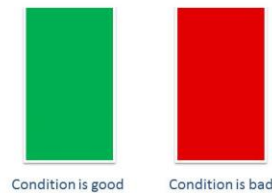
The experiment consists of two parts, each with 8 rounds. One of the rounds of the two parts will be randomly selected at the end of the experiment, and you will be paid your earnings of that round. You will also receive €5 as a fixed show up fee.

[Page 2]

What follows is a description of the tasks for both roles for the first 8 rounds of the experiment. Your decisions in this part have no impact on the procedures in part 2.

General: In each round, every seller is paired with one buyer. Each seller has an imaginary product for sale. This product can be in good condition (with no defects) or in bad condition (with some defects). The seller knows whether the product's condition is good or bad, but this is not shown to the buyer. The buyer's task is to guess whether the product is in good or bad condition, and the buyer earns more money if the guess is correct. The seller earns more money if the buyer guesses that the product is in good condition, no matter in which condition the product is. The exact earnings will be shown on the next page.

Determination of the product's condition: Whether the product is in good condition or bad condition is determined randomly by a card draw. At the start of the round, each seller will get a card. The card can be green or red. If the card is green, the product is in good condition. If the card is red, the product is in bad condition.



Over all 8 rounds, each seller will get at least 4 red cards and 4 green cards. For each seller, the computer will determine randomly in which rounds the card will be green and in which rounds it will be red. Hence, this may differ across different sellers. It is equally likely that the card ends up being green or red in any particular round.

Sellers: *Only the seller will see the card.* The seller must then make a claim about the condition of the product. The seller can claim that the product is in a good condition or in a bad condition by claiming that the card in that round is green or red.

Buyers: The buyer does not observe which card is drawn for the seller. After the seller has made a claim, the buyer makes a guess about the condition of the product.

Chat phase: Before making a guess about the product's condition, there is an opportunity to chat. The chat is free format, meaning that you can chat about everything you wish to talk about. However, you are not allowed to exchange contact information such as telephone numbers, and offensive language or threats are not permitted. You will have two minutes to chat. The time remaining is indicated on the screen.

After the chat phase, buyers will make a guess about the product's condition. We will also ask sellers to indicate how certain they are that the buyer will guess that the product is in good condition. We will ask buyers to indicate how certain they are that the color of the card that the seller drew is green or red.

[Page 3]

Earnings:

The seller earns €30 if the buyer guesses that the product is in good condition, independent of the condition of the product. The seller earns €0 if the buyer guesses that the product is in bad condition.

The buyer earns €30 if his or her guess about the product's condition is correct. The buyer earns €0 if his or her guess about the product's condition is incorrect.

Thus, the seller always earns more if the buyer guesses that the product is in good condition. The buyer always earns more if (s)he makes a correct guess about the condition of the product.

The four possibilities are summarized below:

	Seller	Buyer
1. The seller has a green card (the product's condition is good) and the buyer guesses that the product is in good condition.	€30	€30
2. The seller has a green card (the product's condition is good) and the buyer guesses that the product is in bad condition.	€0	€0
3. The seller has a red card (the product's condition is bad) and the buyer guesses that the product is in good condition.	€30	€0
4. The seller has a red card (the product's condition is bad) and the buyer guesses that the product is in bad condition.	€0	€30

[Page 4]

Matching: At the start of each round, you will be paired with a different participant. Within this part, you will never meet the same participant twice. Sellers will draw a new card at the start of each of the 8 rounds.

Feedback: Sellers will not learn what the buyer guessed. Buyers will not be told what the color of the seller's card was. Note also that we will not make public which round has been selected for payment, to ensure that you cannot identify with whom you were matched in that round, and to ensure that the other participant cannot identify you.

[Page 5]

Your Role: Throughout the experiment you will be a **seller** [/buyer]

Summary: Here is a summary of the steps in each round:

1. The computer randomly draws a card for the seller, determining the condition of the product
2. The seller observes the color of the card and makes a claim to the buyer
3. The buyer and seller have two minutes to chat
4. The buyer makes a guess about the product's condition
5. You will go to the next round and will be paired with a different participant

You can go back to other pages of the instructions by using the menu at the top.

If you have any questions, raise your hand and wait for the experimenter.

Once you are done reading the instructions, press ready to continue. At that point, you will no

longer be able to go back to the instructions on the screen, but you will receive a hard copy with a summary of the instructions on your desk.

[The following instructions were provided after completion of Part 1].

Instructions Part 2

What follows is a description of the tasks for both roles for the next (and last) set of 8 rounds of the experiment.

The second part proceeds in a very similar way as the first part. The only difference is that the seller now earns €18 if the buyer guesses that the product is in good condition, and €12 if the buyer guesses that the product is in bad condition. The earnings for the buyer remain the same as in part 1.¹³

The four possibilities are summarized below:

	Seller	Buyer
1. The seller has a green card (the product's condition is good) and the buyer guesses that the product is in good condition.	€18	€30
2. The seller has a green card (the product's condition is good) and the buyer guesses that the product is in bad condition.	€12	€0
3. The seller has a red card (the product's condition is bad) and the buyer guesses that the product is in good condition.	€18	€0
4. The seller has a red card (the product's condition is bad) and the buyer guesses that the product is in bad condition.	€12	€30

If you have any questions, raise your hand and wait for the experimenter.

¹³ In the Low-Low treatment, this paragraph read: "The second part proceeds in the same way as the first part. To remind you, the possible payoffs are summarized below."

Questionnaire

What is your age?

What is your gender? (Male/ Female)

At which faculty are you a student?

What is your nationality?

What is your native language?

How good do you think you are in general at detecting lies? (7-point scale, from 1 “not good at all” to 7 “very good”)

How good do you think you are in general at deceiving others? ? (7-point scale, from 1 “not good at all” to 7 “very good”)

How often do you play bluffing games like poker? (5-point scale, from 1 “never” to 5 “very often”)

Did you ever take the course Experimental Economics? (no, yes but I did not participate in a similar experiment, yes and I participated in a similar experiment)

We are interested to learn more about the strategies that people use. Can you briefly describe what kind of strategies you used (e.g., were there any things in particular that you wrote to the buyer?)

Participants were also asked to complete a survey developed by Aquino and Reed (2002) that measures “moral identity.”

References:

Aquino, K., & Reed, I. I. (2002). The self-importance of moral identity. *Journal of personality and social psychology*, 83(6), 1423.

Appendix B: Screenshot

Below is a screenshot of the chat environment in experiment 2.

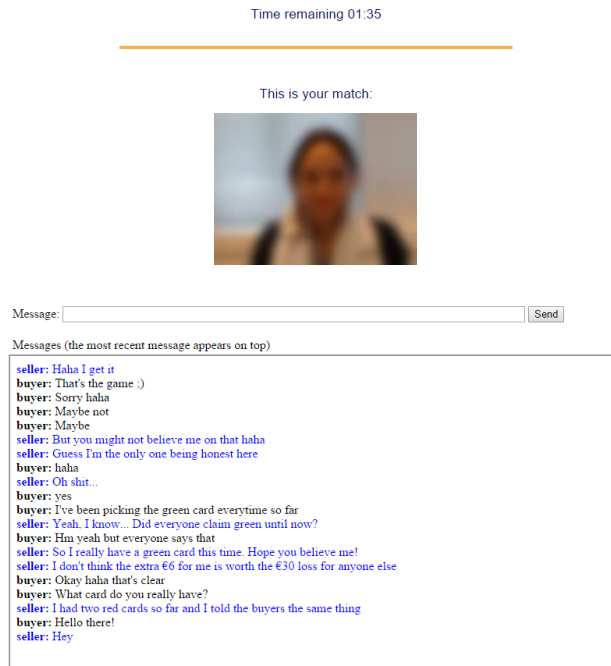


Figure A2-1: Screenshot of the chat environment (Experiment 2). The photo is blurred for privacy reasons. The person on the photo is not related to the chat conversation shown in the screenshot. The chat reads from the bottom to the top (the last message always appears on top).